Available online at www.sciencedirect.com

ScienceDirect

Transportation Research Procedia 5 (2015) 30 – 43

Transportation Research Procedia

www.elsevier.com/locate/procedia

SIDT Scientific Seminar 2013

Individual behavioural models for personal transit pre-trip planners

Agostino Nuzzolo [a], Umberto Crisalli [a] *, Antonio Comi [a], Luca Rosati [b]

[a]Department of Enterprise Engineering, Tor Vergata University of Rome, via del Politecnico 1, Rome 00133, Italy

[b] Department of Civil Engineering and Computer Science Engineering, Tor Vergata University of Rome, via del Politecnico 1, Rome 00133, Italy

Abstract

This paper presents the results of an in-progress research project aiming to define an advanced trip planner for transit networks. Starting from the description of user needs and logical architecture of the trip planner, the paper describes the module to support the user with pre-trip information based on his/her personal preferences. In particular the theoretical aspects of the individual, instead of user group (aggregate), path choice modelling used to support path choice set individuation, path utility calculation and user preference learning process are defined. The theoretical framework has been applied and tested through some experiments carried out on the public transport network of the metropolitan area of Rome.

Peer-review under responsibility of the Società Italiana dei Docenti di Trasporti (SIDT).

Keywords:
ATIS; transit trip planner; personalized advice; personalized pre-trip information; user preference learning process; incremental learning

1. Introduction

For many cities, making public transit more attractive, faster and more efficient is a key rule for increasing modal shift from private to public transport. Transit agencies continuously explore new ways to maintain existing passengers and to attract new ones. Providing static and real-time transit information using new strategies is becoming a priority in many transit agencies around the world. These strategies include the use of new technologies

* Corresponding author. Tel.: 39-06-72597053; fax: +39-06-72597053 .
E-mail address: crisalli@ing. uniroma2.it

Peer-review under responsibility of the Società Italiana dei Docenti di Trasporti (SIDT).
doi:10.1016/j.trpro.2015.01.015

such as internet and wireless mobile to disseminate traveller information on transit services to make them more user-friendly (TRB, 2003; Rizos, 2010).

In recent years, the evolution of Intelligent Transport Systems (ITS) has allowed the development of support systems that help transit agencies in the arduous tasks of planning and managing transit networks increasingly complex and integrated to satisfy the growing mobility needs of users. Today, transit users expect to have a quick and always available comprehensive information about multiple modes (including traffic information) coming in one place or from one source, and on a variety of media.

The more general example of traveller information systems, which can be defined under the umbrella of ITS, are the Advanced Travel Information Systems (ATIS) from which, the branch dedicated to transit are represented by the Advanced Transit Traveller Information Systems (ATTIS). An ATTIS is a specialised information system able to access, organize, summarize, process and display information to help users to plan their individual trips. It may cover a single mode of transport (e.g. bus) or many transport modes for an intermodal journey (e.g. car, bus, metro, rail, including different transit services operated by different transit companies). ATTIS overcome the historical function of route planners, which has traditionally covered just the "route", showing a path by which it is possible to travel between origin and destination at any time; in contrast ATTIS take also into account the transit timetable of services (scheduled, historical or real-time) that run over the network only at certain times, and so the time of travel is relevant when computing a trip.

Advisory tools, such as trip planners, can be considered part of ATTIS. They allow user to easily access to organized information in order to compare the different alternatives for a rational choice of the transport mode (Caulfield and O'Mahony, 2007). As the main reason which leads to reject transit as travel mode is the uncertainty about routes and timetable, transit trip planners have to provide accurate pre-trip information to reduce this uncertainty. The state-of-the-art presents many papers (Adbel-Aty, 2001; Kenyon and Lyons, 2003; Grotenhuis et al., 2007; Tang and Thakuriah 2011; Zhang et al., 2011) that demonstrate the added value of ITS (Intelligent Transport Systems) to improve transit ridership.

Trip planners usually refer to applications through which user can self-access the ATTIS using an user-friendly interface on computers or mobile phones connected to internet, which enable a path search engine for a given origin and destination able to find the best travel alternatives (paths) on transport networks including their relevant information, such as travel time, monetary cost, estimated departure and arrival time, service characteristics, alerts, disruptions. Path search may be optimised on different criteria (e.g. fastest, shortest, least changes, cheapest) and may be constrained to a certain time (e.g. a desired arrival time at destination or a desired departure time from the origin) or to avoid certain waypoints.

ATTIS and relative traveller tools mainly depend on two important characteristics (Levinson, 2003; Toledo and Beinhaker, 2006): data used as the inputs to the computation (static or real-time) and time in which the user receives information, i.e. pre-trip and/or en-route (TRB, 2003; Rizos, 2010).

Static traveller information systems are based on information about the transit network, such as distances, commercial speeds, classification of road facility types and scheduled timetable of all transit services that are defined as planned and that do not consider fluctuations due to unplanned events. An evolution of static information is represented by the updating of such information with the historical one derived from a historical database that represents past traffic conditions. Even if the use of historical information captures the average prevailing transit performances and operations, it does not capture the fluctuations in demand and supply generated by unexpected events.

Real-time traveller information systems allow us to overcome the above limits, as they are based on real-time estimates of current travel times. These estimates are based on traffic information collected by different sources, such as AVL (Automatic Vehicle Location), FPD (Floating Passenger Data), and in the new era of social media by crowdsourcing, too. In contrast with static information, the data analysis is done in real-time, which sets higher computational requirements on models and algorithms for path search. In fact, the real-time computation is more elaborate and involves not only estimation of current conditions but also the use of these estimates as inputs to methods that predict short-term future traffic conditions and non recurrent congestion. In such cases the accuracy of ATTIS strongly depends on the reliability of forecasting methods (Ben-Elia et al., 2013). The information provided at this level captures the effect of time-varying demand and supply due to planned (e.g. maintenance work) and unexpected events (e.g. incidents or disruptions). The generation of information may use statistical or data-based algorithms, as well as transportation models. Statistical methods use real-time traffic measurements to update historical information of travel times in future time periods (e.g. Kalman filter methods or neural networks) but they do not incorporate the effect of user behaviour to the information provided, which can be considered by using model-based approaches, that use simulation-based traffic flows and demand models. Today, the main problem of real-time systems besides in the need of computationally powerful systems and in the significant computational times.

Pre-trip information is the one the user accesses before starting his/her trip. It usually concerns paths, schedules, arrival times, delays, interchanges and other useful multimodal information to travel between origin and destination. Pre-trip information systems allow passengers to plan a door-to-door (or station-to-station) trip using one or more transit services. Pre-trip information may be provided in several different ways, such as websites, and it does not necessitate to track people.

The importance of providing transit information does not stop once the user start his/her trip (pre-trip information) but plays a key role during the trip in keeping the user informed about the status of transit operations, reducing the anxiety, and directing him/her to the right stops, platforms, and bays. Therefore, the *en-route information* describes the real-time transit operations and includes updates on delays, incidents, and service diversions along transit routes, as well as estimated vehicle arrival and departure times for stops along the routes to be used for user-tailored path advice. En-route information may be provided through various media and communication technologies. The most comprehensive way is based on a two-way communication between the ATTIS and the user, and the capability to track the user so that path advice can be user-tailored and provided everywhere on the network.

The use of real-time information including updates about transit services requires specific technological infrastructure that include AVL (Automatic Vehicle Location) systems, communication systems, prediction algorithms and, last but not least, information dissemination media. In particular, the key of success of ATTIS and relative tools (e.g. trip planners) besides in the used media to disseminate the information, which are briefly described below considering the above two important characteristics (static vs. dynamic data and pre-trip vs. en-route information). In recent years, the evolution of information technology and telematics have greatly impacted in the information broadcasting. Currently, transit agencies are using a variety of media to better inform their passengers about services. These media include mobile phones, tablets, DMSs (Dynamic Message Signs), kiosks, Internet and social media, like Facebook and Twitter. Using wireless communications travellers can receive information anywhere and anytime through smartphones and tablets, and through DMSs at stops. The introduction of the Internet and kiosks to provide detailed real-time traveller information allows users to access their personal trip information, while DMSs at stops play the role of disseminating shared information, as the expected arrival times of all transit vehicles at stops.

Information dissemination media can provide personal information (i.e. information provided to the individual traveller), as well as SMS and e-mail, through:

- interactive wayside devices;
- personal communication devices;
- internet and social media.

Interactive wayside devices. People without personal communication devices can use interactive wayside devices such as kiosks, which are located at major bus centres, train stations, and other public places, i.e. hotels, government offices and commercial centres.

Personal communication devices. This category includes traditional wireless devices such as cellular phones and tablets. Wireless communication devices today are spreading because transit agencies can provide high-level customer services at limited costs. Wireless devices are not bounded to the en-route real-time information, but they are also used to provide personal pre-trip information, that in the most advanced TTIS and related tools are implemented in the ATATs (Advanced Transit Trip Advisors) apps.

Internet and social media. Internet and social media represent the state-of-the-art of the information dissemination media, especially for the new frontier of information enabled by the two-way communication opportunities and by the crowdsourcing. In particular, travel tools (e.g. trip planners) based on internet and social media allow a further improvement in the real-time information updating on schedules, vehicle arrivals, paths, and other relevant personal trip information. Moreover, the use of crowdsourcing can improve the accuracy of real-time information provided.

In the sphere of the above described ATTIS framework, this paper focuses on Advanced Traveller Advisory Tools (ATATs) able to assist users in multimodal networks, and to suggest to each user the best path set according to user personal preferences. In order to find the best personal paths, the presented ATAT uses the Random Utility Theory framework (Ben-Akiva and Lerman, 1985) to assign to each path alternative an estimation of the path utility perceived by the user, through which a ranking of such paths is suggested to the user for the choice of the best path that is then presented in details. As it will be reported in the following, this individual perceived utility is function of several attributes (e.g. travel time, walking time, etc.) and their effects on utility is taken into account by model parameters that depend on user personal preferences. Thus, the parameters of an individual path choice model have to be estimated.

In the following, section 2 reports a concise literature review, while section 3 illustrates the user needs and the trip planner logical architecture. Section 4 presents the modelling framework specified to provide personalized information, while section 5 describes the investigation of the theoretical aspects through some experimental results carried out by two test cases on the transit network of the metropolitan area of Rome (Italy). Finally, section 6 reports some conclusions and the future developments of this research.

2. Literature review

Except for laboratory choice experiments in psychology (Thurstone, 1927), it is rare to see discrete choice models estimated for single people. After Chapman (1984), there was little work on ways to measure and model individual choices in survey applications until recently (Frischknecht et al., 2011). In fact, demand models are traditionally used to simulate the average number of trips of given characteristics undertaken by homogeneous user groups and it is not easy to obtain large choice samples of single decision-maker (it is easier to have choice samples from many decision-makers). Therefore, instead of disaggregate or individual behavioural models, user groups (homogeneous with respect to their socio-economic attributes, parameters and the functional form of the models) are used and aggregate behavioural models have been developed. Different types of aggregate models have been proposed, but their performances could seem limited if they are applied to single individuals, because of the variations in taste or preferences among different users.

Some aspects of the individual model development are here discussed. The errors that could be done by using aggregate models are pointed out through the comparison of individual models estimated for users belonging to homogenous groups. Besides, the problem related to the estimation of discrete choice model parameters with repeated observations for each respondent are also to be investigated because, as largely detailed in the literature, it could cause an obvious correlation of disturbances, or heterogeneity (e.g. the parameters can also vary for the same user according to the travel scope), which refers to variations in unobserved contributing factors across behavioural units. If behavioural differences are largely due to unobserved factors, and if unobserved factors are correlated with the measured explanatory variables, then estimates of model coefficients will be biased, when heterogeneity and correlation are not properly treated.

In recent years, researchers have sought to develop ATATs that support user travel and both theoretical and practical solutions were proposed (Arentze, 2013 and Nuzzolo et al., 2014a), but further studies should be still carried out, in particular on the new opportunities given by crowdsourcing. With crowdsourcing systems, as for example Waze (2013) or Moovit (TranzMate, 2013), the choice model can to be user-tailored by learning user’s interests and preferences from usage data. To obtain the individual utility parameters that take into account the individual preferences, a two-step user learning procedure can be used. The first step serves to initialize the path utility function parameters of a new user. The second step uses revealed choices during the use of the traveller tool in order to update the initial parameter estimations.

ATTISs have to define travel alternatives (paths) both in space (among stops) and in time (in relation to the user desired arrival/departure time and/or the arrival/departure time of transit vehicles at stops). In fact, transit networks are usually characterized by different boarding stops and available runs for the same Origin-Destination (OD) pair and target time (e.g. desired arrival time). For this reason, advanced transit trip planners have to use a path modelling approach that explicitly takes into account the above space-time features of path alternatives, such as the schedule-based approach recalled in the section 4. Moreover, ATATs have to provide path alternatives according to single user preferences, for which different path alternatives are suggested to different users according to their different weights on path attributes (e.g. walking time, waiting time, on board time, transfer time and so on). This aspect highlights the frontier of advanced transit pre-trip planners, which have to provide personalized real-time information by using learning process mechanisms able to track profiled travellers and to give them information according to their personal travel habits (Arentze, 2013; Nuzzolo et al., 2014a).

A tool of this type has been developed at the Transport Centre of the "Tor Vergata" University of Rome (Nuzzolo et al., 2014a). It includes a pre-trip planner for multi-service transit networks to suggest a set of alternative paths to each user. These paths are evaluated as the best for a given user according to individual utility parameters, and are hence proposed in their utility order. To obtain the individual pre-trip path choice model and the utility parameters that take into account the individual preferences, a two-step procedure was analysed. The first step initializes the path utility function parameters of a new user on the basis of Stated Preference (SP) Interviews. In this step, each respondent is provided several times with some alternative paths; among them, he/she has to choose the perceived best one. The second step, which updates the initial parameter estimation, uses set of revealed choices during the tool use in a sequential Maximum Likelihood parameter estimation. In this context, the paper presents the main theoretical and experimental results obtained up to now. In particular, the comparison of utility functions obtained for different users is presented and, even if homogeneous with respect to user attributes, parameters and functional form, the importance to use individual instead of aggregate models is hence pointed out. Some analyses on learning process have also been carried out and some preliminary conclusions on the initialization and updating process are presented.

3. User needs and logical architecture

This section describes the user needs and the logical architecture (ARTIST, 2003; Caulfield and O'Mahony, 2007) of the trip planner defined to support transit users with pre-trip personalized information. The main results the trip planner has to provide are:

- transit path alternatives, eventually ranked, according to some user preferences (e.g. maximum number of transfers, maximum walking distance and so on);
- path attributes (departure time, walking times and distances, waiting time at stops, on-board travel time, crowding, transfer number and times, and so on) based on the real transit system operations.

In order to support transit users with personalized pre-trip information, Figure 1 illustrates the logical architecture of the trip planner.

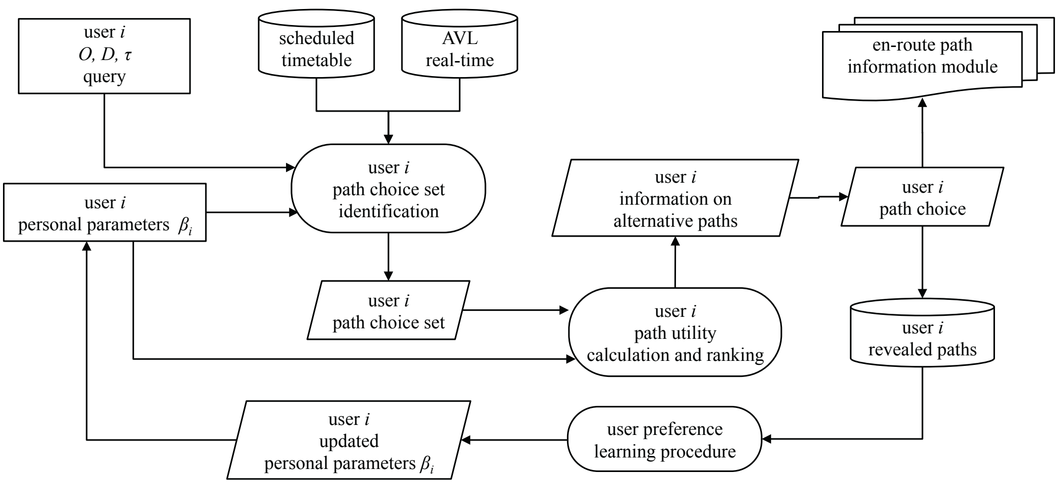

Figure 1 – Logical architecture of the pre-trip personalized information module

The pre-trip module is enabled by a query of the registered user *i*, who is logged into the system. At time τ user *i* requests a support to travel on the OD pair *od* with a desired arrival time τ_{Ai}, then the system identifies the path choice set of user *i* and ranks path alternatives according to his/her personal preferences and to the current information on the multiservice transit network operations (i.e. scheduled timetable and real-time data). Such activities are carried out by using the path choice set identification and ranking procedure described in section 4.

The ranking of alternative paths to the user *i*, is carried out using the modelling framework of the Random Utility Theory (Ben-Akiva and Lerman, 1985), in which personal utility parameters β_i of user *i* are used to estimate path utility for all paths of the path choice set of user *i* (see section 4).

The path choice of user *i* is then added to the personal database of revealed preferences of user *i*, which is utilised to update the personal parameters β_i by using a user preference learning procedure.

The path choice of user *i* also represents the main input of the en-route path information module, aiming to support user *i* during the trip (Figure 1). Further details on the overall (both pre-trip and en-route) logical architecture of the trip planner can be found in Nuzzolo et al. (2014b).

4. Transit path choice modelling

Path choice models can be used to calculate the utility associated with each path *k* belonging to a set of possible alternative paths defined according to traveller characteristics, transport system performance (e.g. travel times and costs) and traveller behavioural assumptions.

As the arrival/departure time coordination of runs at interchanges enables path alternatives, a path modelling approach able to explicitly consider the space-time features of both demand and supply, like the schedule-based approach (Nuzzolo and Crisalli, 2004), has to be used. The schedule-based approach (Nuzzolo et al., 2001) refers to services in terms of runs (vehicles) using the real vehicle arrival/departure times, and hence all the values of level of service attributes, evaluated at time in which users make their choices, can be explicitly considered. It allows us to take into account the evolution in time of both supply and demand, as well as run loads and level of service attributes, by an explicit treatment in modelling the:

- demand temporal segmentation, to consider user desired departure or arrival times;
- supply, to represent each run of transit services with its departure/arrival times at stops;
- path choice, to take into account the attribute time-dependencies.

On the demand side, the temporal segmentation of the demand in relation considers both a Desired Arrival Time at destination (DAT), represented by τ_{Ai}, and a Desired Departure Time from origin (DDT), represented by τ_{Di} . It is obtained by dividing the generic day *t* in *n* elementary time intervals of $\Delta\tau$ width (e.g. $\Delta\tau_i=1$ minute), to which the user target times τ_{TTi} (DAT and/or DDT) requested by the user are associated to.

On the supply side, path attributes and on-board loads for each run of transit services can be calculated using a diachronic network (Nuzzolo et al., 2001), whose diachronic graph Ω consists of three different sub-graphs made of nodes with explicit time coordinates (Figure 2). The diachronic graph Ω consists of:

- a service sub-graph, in which each run of each line is defined both in space, through stops, and in time, according to arrival/departure times at stops;
- a temporal centroids sub-graph, in which each node represents both temporal centroids, in order to simulate the space-time characteristics of trips, and user arrival/departure times;
- an access-egress sub-graph, which allows the connection between centroids and stops, and stops between them.

Regarding path choice, a path *k* on a transit network is the space-time sequence of transport infrastructures and services used by an user travelling from an origin *o* at a given origin departure time τ_{Di} to a destination *d* with the relative arrival time at destination τ_d, which also includes access stop *s* with relative arrival time at stop τ_{Dis}, line and run with run departure time τ_r from access stop (or sequence of lines and runs including the relative stop interchanges) and egress stop *s'* allowing the user to reach his/her destination. It can be defined through a sequence of links of the Ω graph. As the diachronic network is represented through a graph, we can apply the entire traditional network algorithms (e.g. least cost) to efficiently obtain path search and space-time attributes. Furthermore, the explicit representation of single runs with their characteristics (e.g. capacity) allows us to explicitly treat congestion through vehicle capacity constraints (Nuzzolo et al., 2012).

Path choice model are the core of the proposed trip planner, and for this reason the following sections will describe the main features of the presented modelling framework deepening on path choice set generation, path utility calculation and model calibration on the basis of single-user personal preferences.

In order to test the trip planner, the presented theoretical modelling framework is supported by the results of some experimental evidences carried out in the metropolitan area of Rome (Italy), which is served by a multiservice transit network (urban bus, tram and metro, regional railways and buses) operated by different companies with an integrated fare structure.

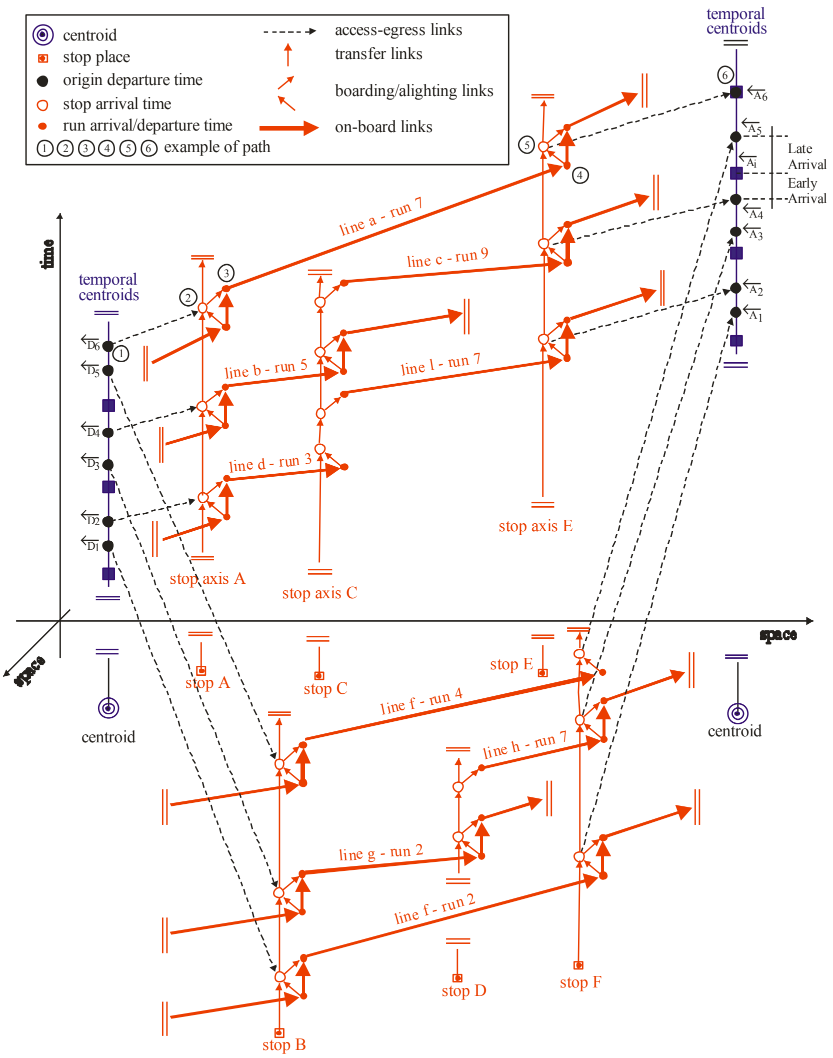

Figure 2 – Example of path choice alternatives on a diachronic graph

4.1. The pre-trip path choice set generation

For user *i* travelling on the OD pair *od* at time τ_{TTi}, the path choice set can be defined through a selective approach based on a set of rules that allows us to define feasible paths according to traveller characteristics, transport system performance (e.g., travel times and costs) and traveller behavioural assumptions. The rules to reduce the potentially high number of path alternatives can be heuristically calibrated on the basis of a sample of observed choices (maximum coverage factor method).

For example, one of the the path choice sets used in the experimental analysis of section 5 is here described. Given user *i* travelling from Frascati (a town in the suburbs of Rome) to Piazza Sempione (i.e. centre of Rome) with the desired arrival time at 9:30 am, the four different paths schematically pictured in Figure 3 can be defined. This OD pair is characterised by a distance on the road network of about 25 km and by an average travel time of about 2 hours. The above four path alternatives differ in terms of travel time, waiting and transfer times, modes to be used (train, metro or bus) and early/late arrival time.

Figure 3 – Example of path choice alternatives

4.2. The single-user path utility model

The utility $V^{\tau}_{od,\tau_{TT_i}}[k]$ that user *i* at time τ associates to the path *k* identified by departure time τ_{Di}, access stop *s* and run *r*, can be expressed as:

$$V^{\tau}_{OD,\tau_{TT_i}}[k] = \beta_{ED}\cdot ED_k + \beta_{AE}\cdot AE_k + \\ + \sum_m [\beta_{TW,m}\cdot TW_{m,k} + \beta_{OB,m}\cdot OB_{m,k} + \beta_{CFW,m}\cdot CFW_{m,k} + \beta_{MP,m}\cdot MP_{m,k} + \beta_{NT,m}\cdot NT_{m,k}] \tag{1}$$

where ED_k is the Early or Late arrival time (i.e. the difference between the desired and the actual arrival time at destination) using path *k*; AE_k is the sum of access and egress times on path k; $TW_{m,k}$ is the waiting time spent for boarding runs of the transit service *m* (train, metro, tram, bus) belonging to path k; $OB_{m,k}$ is the on-board time spent on the transit service *m* belonging to path *k*; $CFW_{m,k}$ is the average on-board crowding degree on runs of transit service *m* belonging to path *k*; $MP_{m,k}$ is a preference attribute for transit service *m* on path *k* (e.g. expressed as a function of the travel distance on transit service *m* w.r.t. the total distance on path *k*); $NT_{m,k}$ is the number of transfers on transit service m belonging to path *k*; β_i are the model parameters.

The reader should note that in the sphere of a trip planner, eqn (1) is dynamically estimated at time τ on the basis of time-dependent choice sets and attributes for each path alternative. Furthermore, most of attributes of eqn(1) are provided by ATTIS on the basis of the real-time information on transit operations.

Once calculated path utilities, the ranking of path alternatives can be carried out considering such utilities. Then, it is possible to select the one/ones (e.g. one or more) the trip planner will suggest to the user. Moreover, aiming to provide personalised information, a personal (individual) set of model parameters β_i has to be used in (1). These parameters are estimated considering the single-user preferences/attitudes as follows.

4.3. The estimation of individual pre-trip path utility parameters

The calibration of individual path choice parameters implies that the model functional form could be the same for different users, as the same could be the values of attributes considered in path choice, but different sets of parameters (i.e. different for each registered user and travel purpose) have to be considered and calibrated according to user travel preferences.

The estimation of individual coefficients β_i can be performed using the information collected from a sample of observations. Given a sample of *N* observations of a single user *i*, the problem is to find the estimates of coefficients β_i though which the travel planner is able to rank the alternatives and suggest the user best perceived paths. For this reason, some aspects of the individual model development should be considered, as here presented.

The estimation of discrete choice model parameters with repeated observations for each respondent gives rise to an obvious correlation of disturbances, or heterogeneity (e.g. the parameters can also vary for the same user according to travel purpose), which refers to variations in unobserved contributing factors across behavioural units. If behavioural differences are largely due to unobserved factors, and if unobserved factors are correlated with the measured explanatory variables, then estimates of model coefficients will be biased if heterogeneity and correlation are not properly treated. The problem may be more pronounced in repeated measurement data since unobserved factors may be invariant across these repeated measurements. As the panel data contains multiple observations of the same user, the assumption of independence between choices for the same user may not be appropriate. For this reason, growing interest in the representation of unexplained heterogeneity in choice data arose by using random coefficients models, such as Mixed Multinomial Logit (Hess and Rose, 2009). Moreover, different model specifications could be used (e.g. nested logit models, error components logit model) to consider that the paths could be not completely independent due to possible overlaps.

5. Experimental evidences

In order to investigate the individual path choice model, preliminary tests based on some experimental evidences were carried out using simple multinomial logit models (MNL) for their well-known easy-to-apply advantages. Such models were performed by considering the working-day journeys on two different OD pairs.

The first considers a sub-urban trip from Frascati (a town near Rome) to the centre of Rome (i.e. Piazza Sempione) with the desired arrival time at 9:30 am, whose path alternatives are detailed in Figures 3 and 4.

The second test was carried out for the urban O/D relation, within the city of Rome (from Via Monvisio to Via Tivoli) with a distance of about 9 km and with three transit path alternatives: one including metro interchange and the others using only bus lines, as schematised in Figure 5.

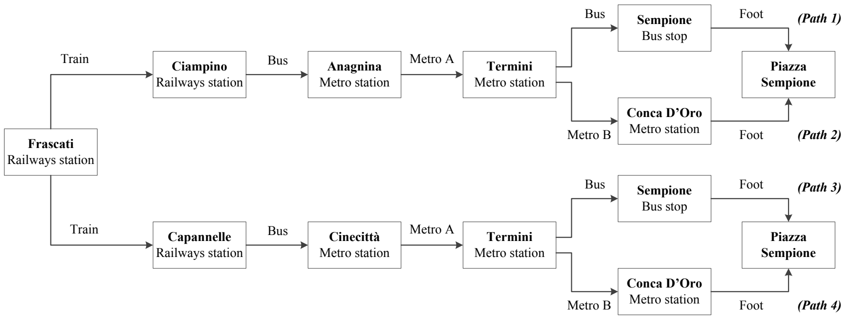

Figure 4 – Path alternatives for the sub-urban OD test case

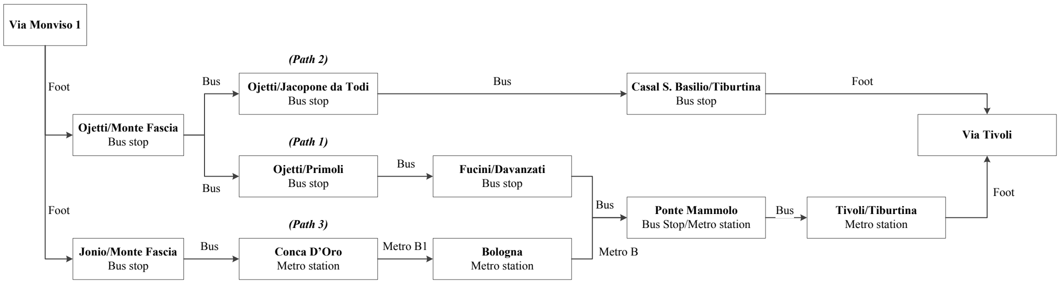

Figure 5 – Path alternatives for the urban OD test case

Path choice models were estimated carrying out a SP survey with 150 sets of four scenarios considering 6 test users (A, B, C, D, E, F) for the sub-urban OD pair and 2 test users (G, H) for the urban one, to which the choices of the preferred paths were asked. The set of 150 scenarios with 4-alternatives was defined by path alternatives randomly extracted from the previous experimented status of the transportation system operating in the study area.

Individual models for different travellers are presented in Tables 1 and 2, showing relevant differences among the preferences of the travellers in both the sub-urban and urban test cases.

Although all the attributes displayed in Equation (1) were tested, only those reported in Tables 1 and 2 were statistically significant with different utility functions specified for the sub-urban and urban test cases, for which the sub-urban case shows a greater articulation of attributes composing the utility function, as expected, due to the greater complexity of the trip (e.g. multimodal and/or multiservice).

Table 1 – MNL model parameter (sub-urban OD test case)

Multinomial Logit (MNL)								
attributes	*units*	*parameters*						
		individual User A	individual User B	individual User C	individual User D	individual User E	individual User F	**Aggregate**
Waiting time (total)	minutes	-0.439 *(-4.06)*	-1.220 *(-5.50)*	-0.073 *(-3.31)*	-0.773 *(-3.65)*	-0.497 *(-5.24)*	-0.105 *(-1.71)*	**-0.219** ***(-11.59)***
On-board time (train)	minutes	-0.078 *(-1.30)*	-0.430 *(-3.12)*				-0.261 *(-4.90)*	**-0.062** ***(-3.96)***
On-board time (metro)	minutes	-0.336 *(-1.64)*	-0.830 *(-2.80)*			-0.516 *(-6.40)*	-0.026 *(-1.14)*	**-0.080** ***(-1.58)***
On-board time (bus)	minutes	-0.426 *(-3.24)*	-0.590 *(-3.31)*	-0.265 *(-6.81)*	-1.100 *(-3.24)*		-0.179 *(-1.73)*	**-0.180** ***(-6.29)***
Early and late arrival time	minutes	-0.291 *(-3.41)*	-0.570 *(-3.27)*	-0.081 *(-2.58)*	-0.074 (-1.67)		-0.169 *(-2.90)*	**-0.036** ***(-2.29)***
ρ^2		0.71	0.67	0.60	0.90	0.79	0.69	**0.51**
%-of-right		84%	81%	64%	95%	88%	90%	**76%**
%-of-right incl. 1st+2nd best		99%	98%	100%	100%	100%	96%	**90%**

(-) = t-st value

Table 2 – MNL model parameter (urban OD test case)

Multinomial Logit (MNL)				
attributes	*units*	*parameters*		
		individual User G	individual User H	**Aggregate**
Total travel time	minutes	-0.199 *(-7.86)*		
Time on foot	minutes		-0.235 *(-5.11)*	**-0.062** ***(-1.78)***
On-board and transfer time	minutes		-0.151 *(-5.24)*	**-0.187** ***(-10.67)***
Metro ASC		-2.558 *(-6.24)*	-1.175 *(2.10)*	**-0.339** ***(-2.31)***
ρ^2		0.85	0.46	**0.41**
%-of-right		75%	72%	**70%**
%-of-right incl. 1st+2nd best		100%	94%	**94%**

(-) = t-st value

Deepening on the results of the sub-urban case in terms of heterogeneity of attributes among individual users, and considering as reference unit the weight of the on-board time by bus (OB_{bus}), Table 1 and Figure 6 show the different perception of waiting time (WT) among users. In fact, except for user B that perceives the waiting time about the double w.r.t. on-board time by bus, all the others perceive that time equal (users A and E) or smaller (users C, D, E);

the use of the aggregate model in such case, which presents a weight of the waiting time larger than one w.r.t. on-board time by bus, seems to overestimates the role of the waiting time for most of the considered users.

The same analysis can be done in terms of early or late arrival time (*ED*), again w.r.t. on-board time by bus, for which we can see that for users that consider this attribute (users A, B, C, F), some of them (A and B) give the same importance w.r.t. on-board time by bus; in any case all give a greater importance with respect to that reproduced from the aggregate model that seems to underestimate this aspect.

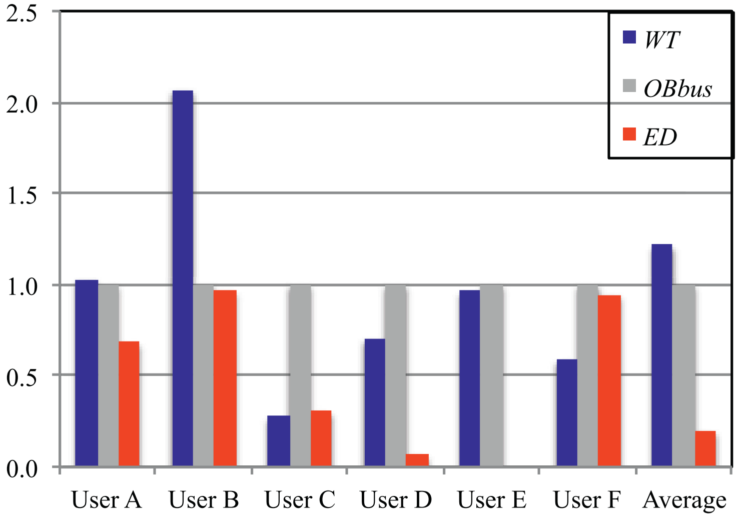

Figure 6 – Perception of attributes for the sub-urban OD test case

For what concerns the results obtained for the urban case, Table 2 shows similar weights for the travel time, even if the reproduction of choices of user H requires a greater articulation of travel time distinguishing in time spent on foot and the rest (on-board and transfers), where the first weighs more than the rest (about 1.5 times). On account of the test case (urban OD pair), for which an high- frequency of services can reasonably be assumed, waiting time was not statistically significant and was not introduced in the utility function.

Moreover, the capabilities of individual models to reproduce user behaviours is shown by the values of ρ^2 and *%-of-right* (calculated both for the choice of best path alternative and for that of the first and the second ones), which always present values better that those obtained for aggregate models.

Aiming to use such models within ATATs, model performances should be mainly considered in terms of *%-of-right* for which, except for one case (user C), Table 1 shows that individual models predict more than the 81% of the user chosen paths. If the analysis is extended to the first and the second best paths, the *%-of-right* becomes greater than the 96%, with top values of 99-100% in most cases (4 of 6).

Tables 1 also shows that if an aggregate Multinomial Logit model is used, the performances fall dramatically. Such differences can be better appreciated in Table 3, which reports the above described individual model performances in reproducing user choice for each individual user, compared with that obtained by using the aggregate model. The analysis of values of Table 5 clearly represents a further proof, which confirms the need to apply individual models in order to suggest paths according user personal preferences.

The reader should note that the same phenomenon can be highlighted in the urban case (Table 2 and 4), although with limited differences in heterogeneity reasonably due the reduced size of the user sample w.r.t. to the sub-urban case.

Table 3 – Model performance for the sub-urban OD test case

	%-of-right on best path choices			
	individual model		aggregate model	
	1st	1st+2nd	1st	1st+2nd
User A	84	99	83	98
User B	81	98	51	77
User C	64	100	65	89
User D	95	100	91	100
User E	88	100	79	95
User F	90	96	85	96

Table 4 – Model performance for the urban OD test case

	%-of-right on best path choices			
	individual model		aggregate model	
	1st	1st+2nd	1st	1st+2nd
User G	75	100	74	97
User H	72	94	62	94

6. Conclusions

This paper presented some results of an on-going research aiming to define a trip planner able to give pre-trip personalized information to the user about travel alternatives on a transit network fostered by real-time data. The the theoretical framework is based on a path choice model able to provide transit path alternatives on the basis of user personal travel preferences captured by a learning process on user habits. The presented theoretical and experimental analyses investigated the specification of individual models able to suggest the best personal perceived paths.

Model parameters were estimated tracking the choices made by a sample of students travelling for leisure in different contexts (sub-urban and urban), which showed the need of using individual models to suggest the individual preferred paths with a good reliability.

Further developments of this research regard the additional investigation of the path choice modelling and the extension to the en-route personalized information. In particular, advances in path choice modelling concern the investigation of other O/D pairs and larger samples, aiming at exploring user preferences and other model forms (e.g. mixed logit), as well as the user preference learning process, aiming at updating model parameters to improve the accuracy of model prediction and user compliance.

References

Adbel-Aty, M.A. (2001). Using ordered probit modelling to study the effect of ATIS on transit ridership. *Transportation Research Part C*, 9, Elsevier, 265-277.

Arentze, T.A. (2013). Adaptive, personalized travel information systems: A Bayesian method to learn users' personal preferences in multi-modal transport networks. In *Proceedings of Transportation Research Board Annual Meeting*, Washington, USA.

ARTIST (2003). ARchitettura Telematica Italiana per il Sistema dei Trasporti Ministero delle Infrastrutture e dei trasporti, Retrieved August 5, 2013, from *http://www.its-artist.rupa.it*

Ben-Akiva, M., & Lerman, S. (1985). *Discrete choice analysis*. MIT Press, Cambridge, Boston, USA, 1985.

Ben-Elia, E., Di Pace, R., Bifulco, G.N., & Shiftan, Y. (2013) The impact of travel information's accuracy on route-choice, *Transportation Research Part C*, 26, 146-159.

Caulfield, B., & O'Mahony, M. (2007). An examination of the Public Transport Information requirements of users. *IEEE Transactions on Intelligent Transportation System*, 8(1).

Chapman, R.G. (1984). An Approach to Estimating Logit Models of a Single Decision Maker's Choice Behavior. In *Advances in Consumer Research*, 11, T.C. Kinnear (Ed.), 656-661.

Frischknecht, B., Eckert, C., & Louviere, J. (2011). Simple ways to estimate choice models for single consumers. *Centre for the Study of Choice (CenSoC), Working Paper Series, No. 11-006*, University of Technology of Sydney.

Grotenhuis, J.W., Wiegmans, B.W., & Rietveld, P. (2007). The desired quality of integrated multimodal travel information in public transport: Customer needs for time and effort savings. *Transport Policy* 14(1), Elsevier, 27–38.

Hensher, D.A., & Greene, W.H. (2003). The mixed logit model: the state of practice. *Transportation*, 30, Kluwer Academic Publishers, 133-176.

Hess, S., & Rose, J.M. (2009). Allowing for intra-respondent variations in coefficients estimated on repeated choice data, *Transportation Research Part B*, 43, Elsevier, 708-719.

Kenyon, S., & Lyons, G. (2003). The value of integrated multimodal traveller information and its potential contribution to modal change. *Transportation Research Part F*, 6, Elsevier, 1–21.

Lancsar, E., & Louviere, J. (2008). Estimating individual level discrete choice models and welfare measures using best worst choice experiments and sequential best worst MNL. *Centre for the Study of Choice (CenSoC)*, University of Technology of Sydney.

Levinson, D. (2003). The value of advanced traveler information systems for route choice, *Transportation Research Part C*, 11, 75-87.

Molin, E.J.E., & Arentze, T.A. (2013). Travelers' preferences in multimodal networks: design and results of a comprehensive series of choice experiments. In *Proceedings of Transportation Research Board Annual Meeting*, Washington, USA.

Nuzzolo, A., & Crisalli, U. (2004). The schedule-based approach in dynamic transit modelling: a general overview. In N.H.M. Wilson and A. Nuzzolo (Ed.), *Schedule-Based Dynamic Transit Modeling. Theory and Applications* (pp.1-24). Kluwer.

Nuzzolo, A., Crisalli, U., & Rosati, L. (2012). A schedule-based assignment model with explicit capacity constraints for congested transit networks. *Transportation Research Part C*, Elsevier, DOI: 10.1016/j.trc.2011.02.007.

Nuzzolo, A., Russo, F., & Crisalli, U. (2001). A doubly dynamic schedule-based assignment model for transit networks. *Transportation Science*, 35, 268-285.

Nuzzolo, A., Crisalli, U., Comi, A., & Rosati, L. (2014a). Advanced Trip Planners for Transit Networks: Some Theoretical and Experimental Aspects of Pre-Trip Path Choice Modeling. In *Computer-based Modelling and Optimization in Transportation*, J.F. de Sousa and R. Rossi (eds), Advances in Intelligent Systems and Computing 262, Chapter 30, Springer, DOI: 10.1007/978-3-319-04630-3_30

Nuzzolo, A., Crisalli, U., Comi, A., & Rosati, L. (2014b). Advanced Trip Planners for Transit Networks: Some Theoretical and Experimental Aspects of Pre-Trip Path Choice Modeling. In *Proceedings of 11th Congreso de Ingeniería del Transporte (CIT 2014)*, Procedia Social and Behavioral Sciences, Elsevier, Forthcoming.

Rizos, A.C. (2010). Implementation of Advanced Transit Traveler Information Systems in the United States and Canada: practice and prospects, *Thesis for the degree of bachelor of Science in Planning and Master in City Planning*, Massachusetts Institute of Technology.

Tang, L., & Thakuriah, P. (2011). Will the Psychological Effects of Real-time transit information systems Lead to Ridership Gain?. In *Proceedings of Transportation Research Board Annual Meeting*, Washington, USA.

Thurstone, L.L. (1927). A Law of Comparative Judgement. *Psychological Review* 34, 273-286.

Toledo, T., & Beinhaker, R. (2006). Evaluation of the Potential Benefits of Advanced Traveler Information Systems, *Journal of intelligent transportation systems*, 10 (4), 173-183.

Transportation Research Board (2003). Strategies for Improved Traveler Information, *Transit Cooperative Research Program (TCRP)*, Report 92.

TranzMate (2013). Moovit. Retrieved August 5, 2013, from *http://www.moovitapp.com*

Waze (2013). Waze. Retrieved August 5, 2013, from *http://www.waze.com*

Zhang, L., Li, J., Zhou, K., Gupta, S.D., Li, M., Zhang, W.B., Miller, M.A., & Misener, J.A. (2011). Design and implementation of a traveller information tool with integrated real-time transit information and multi-modal trip planning. In *Proceedings of Transportation Research Board Annual Meeting*, Washington, USA.